A Collective Motion Description of the 3_{10}-/α-Helix Transition: Implications for a Natural Reaction Coordinate

Gautam Basu, Akio Kitao, Fumio Hirata, and Nobuhiro Gō*

Contribution from the Department of Chemistry, Faculty of Science, Kyoto University, Kitashirakawa, Sakyo-ku, Kyoto 606, Japan

Received January 31, 1994. Revised Manuscript Received April 26, 1994•

Abstract: The 3_{10}-/α-helical transitions in an eight-residue Aib homopeptide is investigated using the methods of normal mode (NM) and principal component (PC) analysis. Energy minimization is followed by a NM analysis of the two (3_{10}- and α-helix) minimized conformations. From the analysis of the NMs it is shown that the 3_{10}-helix is entropically favored over the α-helix. A 2 ns molecular dynamics (MD) simulation run is then analyzed by the method of PC analysis—a quasi-harmonic method for delineating collective motions of the peptide from the raw MD data. Important information about the mechanism of transition at the molecular level is then extracted by projecting the MD trajectory onto the first few PC axes. The first PC axis, representing about two-thirds of the total fluctuation, is shown to corroborate well with the helical states of the peptide and is a natural pathway in one dimension. Addition of contributions from four more PC axes was sufficient to accurately reproduce the peptide dynamics during the transition. A description of the transition, almost in its entirety, in a conformational subspace spanned by only the first few PCs, is significant, and the general implications for simulations of conformational transitions are discussed.

1. Introduction

In recent years computer simulation studies of peptides[1,2] and proteins[3] have provided significant insight into the dynamic aspect of their structures, and the need to consider the dynamics rather than only the static structures is becoming more and more recognized. Such an approach becomes indispensable in understanding transitions and fluctuations between conformational substates in equilibria. One such conformational equilibrium in short peptides, extensively investigated theoretically and experimentally, is the 3_{10}-/α-helix equilibrium.[4-8]

α-Helices are fairly common structural elements in globular proteins, whereas only short stretches of 3_{10}-helices have been observed to occur in proteins; about 10% of all residues forming helices in general were found to be in short 3_{10}- rather than α-helical conformation in a recent survey.[9] This tendency to form only rare and short 3_{10}-helices in proteins changes dramatically when the amino acid Aib (α-methylalanine or α-aminoisobutyric acid) is incorporated into a peptide. Peptides containing 3–12 Aib residues assume the 3_{10}-helical conformation.[10,11] However, as copolymers of naturally occurring and Aib residues are studied, it is now known that the primary effect of the Aib residues is not to impart an exclusive 3_{10}-helical conformation but to constrain the peptide backbone to a narrow energy minimum containing *both* the α- and the 3_{10}-helices.[4] This poses a subtler question—which one of the two, the 3_{10}- or the α-, will be the favored conformation in such copolymers? Or, will the preferred conformation be a mixed helical kind? This helix–helix equilibrium has been shown to be very delicate and with a subtle change of factors like the peptide length, amino acid composition (percent Aib), amino acid sequence, or the solvent, the equilibrium is shifted one way or the other.[5,8] This becomes especially true when the length of the peptide corresponds to a critical value of about eight amino acids.

Almost all Aib-containing peptides so far studied are synthetic and designed solely for the purpose of studying the helical propensity of the amino acid. However, the recent surge of experimental and theoretical work on Aib containing peptides should not be thought of as only of nonbiological value, providing insights into using Aib in synthetically designed macromolecules. The amino acid Aib frequently occurs in membrane-spanning ion channel forming peptides. Definitely, one of the primary roles that Aib plays in these peptides is to impart a strong helicity. But, to understand the reasons for the preference of Aib over other helix-forming natural amino acids, one may have to study the dynamic aspect of such peptide. As of now there is no clear and precise information about the mechanism of such channels, but the dynamic picture of the helix has been suggested to play a role.[12]

Although dynamic computer simulation studies have been performed on a variety of proteins and peptides, only recently have theoretical studies attempting to address the question of 3_{10}-/α-helix dynamics in Aib-rich peptides started to emerge.[13-15] Also, in several recent MD simulations, performed on proteins[16] and smaller peptides,[17-19] containing no Aib residues, 3_{10}-/α-helix transitions were observed.

* Authors to whom all correspondences should be addressed.

• Abstract published in *Advance ACS Abstracts*, June 1, 1994.

(1) Brooks III, C. L.; Case, D. A. *Chem. Rev.* **1993**, *93*, 2487–2502.
(2) Hermans, J. *Curr. Opin. Struct. Biol.* **1993**, *2*, 270–276.
(3) See, for example, the special issue on Protein Dynamics: *Chem. Phys.* **1991**.
(4) Karle, I.; Balaram, P. *Biochemistry* **1990**, *29*, 6747–6756.
(5) Marshall, G. R.; Hodgkin, E. E.; Langs, D. A.; Smith, G. D.; Zabrock, J.; Leplawy, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 487–491.
(6) Hodgkin, E. E.; Clark, J. D.; Miller, K. R.; Marshall, G. R. *Biopolymers* **1990**, *30*, 533–546.
(7) Marshall, G. R.; Smythe, M. L.; Huston, S. E.; Bindal, R. D. In *13th American Peptide Symposium*; ESCOM Scientific Publishers, Leiden, Holland: 1993.
(8) Basu, G.; Kuki, A. *Biopolymers* **1992**, *32*, 61–71.
(9) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601–619.
(10) Pavone, V.; Di Blassio, B.; Santini, A.; Benedetti, E.; Pedone, C.; Toniolo, C.; Crisma, M. *J. Mol. Biol.* **1990**, *214*, 633–635.
(11) Paterson, Y.; Stimson, E.; Evans, D. J.; Leach, S. J.; Scheraga, H. A. *Int. J. Peptide Protein Res.* **1982**, *20*, 468–480.
(12) Pullman, A. *Chem. Rev.* **1991**, *91*, 793–812.
(13) Smythe, M. L.; Huston, S. E.; Marshall, G. R. *J. Am. Chem. Soc.* **1993**, *115*, 11594-11595.
(14) Huston, S. E.; Marshall, G. R. *Biopolymers* **1994**, *34*, 75–90.
(15) Zhang, L.; Hermans, J., submitted for publication.
(16) Fan, P.; Kominos, D.; Kitchen, D. B.; Levy, R. M.; Baum, *Chem. Phys.* **1991**, *158*, 295–301.
(17) Tirado-Rives, J.; Maxwell, D. S.; Jorgenson, W. L. *J. Am. Chem. Soc.* **1993**, *115*, 11590–11593.
(18) Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1991**, *30*, 3864–3871.
(19) Soman, K. V.; Karimi, A.; Case, D. A. *Biopolymers* **1991**, *31*, 1351–1361.

0002-7863/94/1516-6307$04.50/0 © 1994 American Chemical Society

Further motivation for studying conformational transitions in Aib rich peptides arises from some earlier experimental work reported by one of us.[20] Due to the severe conformational constraint that the Aib residue imparts to the peptide backbone, peptides rich in Aib have been designed and used as idealized and well defined helical environments, where light-induced intramolecular electronic interactions like electron transfer or other nonradiative processes can be studied. The presence of any dynamic redistribution between conformational substates, with a time scale comparable to that of electronic transitions studied, will have direct consequences on the experimental observables. Therefore, it is important to investigate the dynamic aspects of these synthetic peptides from theoretical studies as well. Such experiments can also provide experimental numbers for conformational transition rates and activation energies. Although the rate constant for interconversion between the right-handed and left-handed 3_{10}-helices has been reported,[21] no experimental rate constant is known for the 3_{10}-/α-helix transition.

In this paper we present results of normal mode (NM) analysis and molecular dynamics (MD) simulation results for an eight-residue homopeptide of Aib for which the 3_{10}-/α-helix equilibrium constant was estimated to be close to unity from conformational energy analysis.[22] NM analysis is carried out at two energy minima corresponding to α- and 3_{10}-helices. This analysis, based on the assumption of harmonicity of the conformational energy surface at each minimum, expresses the dynamics of the peptide by a superposition of collective motions called NMs and is suited for evaluating the conformational entropy of each helical form. In the MD simulation no such assumption is made, and the result of our simulation does indeed show the 3_{10}-/α-helix transition as was also observed in earlier computational studies.[13-19]

What is new in our computational study is development and application of a new methodology to identify natural reaction pathways of the helix–helix transition. In the earlier studies end-to-end distance[13,14] or homogeneous concerted change to backbone angles[17] were chosen as reasonable but somewhat arbitrary guesses of the reaction coordinate. In this paper we will derive natural reaction pathways from the trajectory of MD simulation. The method developed for this purpose is based on the concept of "important subspace", which has been recently introduced and used successfully in the X-ray crystallographic refinement method, the NM refinement.[23] The concept of important subspace means that, even though the conformational energy surface for peptides and native proteins are so anharmonic as to have many different local minima, the subspace spanned by a relatively small number of low-frequency NMs calculated at any one of the local minima does not depend much on the minimum where it is calculated, and most of the important conformational events, harmonic or anharmonic, occur in this subspace. According to this concept, helix–helix reaction pathways for Aib homopeptides should also exist in the important subspace. In this paper we seek for the natural reaction pathways in a subspace spanned by a small number of principal components (PCs), not by NMs. PCs are defined as vectors which diagonalize the variance–covariance matrix of deviation of mass-weighted cartesian atomic coordinates from their mean. If such an important subspace, spanned by a small number of NMs, exists for Aib homopeptide that is independent of the minimum where NM vectors are calculated, and if the phenomenon of 3_{10}-/α-helix transition is indeed taking place in such a subspace, then this subspace should also be identifiable as spanned by a small number of PCs. We will show that, by the method based on this principle, natural reaction pathways can be identified in a space spanned by a few PCs, each representing a collective motion in the Aib homopeptides.

2. Materials and Methods

The peptide studied, AIB8, is an eight-residue Aib oligomer blocked at the N-terminus with an acetyl and at the C-terminus by a methyl amide group (Ac-Aib_8-NHMe). Eight-residue long Aib oligomers are known to form the 3_{10}-helical conformation in solution and in the solid state.[24]

Energy minimizations were performed by the program IMPACT[25] with the all-atom AMBER force field. The same force field including the partial charges for the amino acid Aib was used to be consistent with the previous work.[6] The relative dielectric constant was taken equal to the distance in Å. Energy minimizations, performed with a combination of the steepest descent and the conjugate gradient methods, were started with idealized helices (α: $\phi = -55°, \psi = -45°$; 3_{10}: $\phi = -60°, \psi = -30°$).

A 2 ns long vacuum MD simulation with a distance dependent dielectric constant was performed with the program IMPACT.[25] Bond lengths were constrained with the SHAKE algorithm,[26] and no cutoff was used. The temperature was maintained at 298 K by weakly coupling[27] the peptide to a heat bath with a relaxation time of 0.1 ps. The starting conformation was the energy-minimized 3_{10}-helix which was equilibrated for 20 ps before trajectory information was collected at every 0.2 ps.

NM analysis is an established method for studying the dynamics of peptides and proteins within the harmonic approximation.[28-30] In this paper NM analysis in the mass-weighted cartesian coordinate space was carried out using the standard procedure.[31] In the NM approximation the displacement of the mass-weighted position vector of the ith atom from the potential energy minima, ΔX_i ($= m_i^{1/2}$), is given in terms of the NM coordinates, Q_j, as

$$\Delta X_i = \sum_{j=1}^{3N-6} \alpha_{ij} Q_j \tag{1}$$

where m_i and Δx_i are the mass and the displacement of Cartesian coordinates of the ith atom. The matrix α, with elements α_{ij}, transforms the hessian F to a diagonal matrix ω containing the angular NM frequencies squared, and the NM vectors are normalized as

$$\alpha^T \alpha = \mathbf{I} \tag{2}$$

In the present work the hessian of the conformational energy was calculated numerically from the analytically calculated gradient.

The MD trajectory was analyzed by the method of PC analysis. This method[32] is a very useful way to analyze collective motions in a MD simulation. Each conformation in the MD trajectory is represented by a point in the multidimensional conformational space giving rise to a distribution representing the particular MD run. The variance–covariance matrix of this distribution represents fluctuations during the MD run, and the PC axes are defined as those which diagonalize the variance–covariance matrix. Those PC axes that have large variances can then be conveniently identified and studied to understand the collective motion.

The variance–covariance matrix, r, can be written as

$$r_{ij} = \langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle \tag{3}$$

The PC matrix U, whose jth column vector is the jth PC axis, is solved such that

(20) Basu, G.; Kubasik, M.; Anglos, D.; Kuki, A. *J. Phys. Chem.* **1993**, *97*, 3956–3967.
(21) Hummel, R.-F.; Toniolo, C.; Jung, G. *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 1150–1152.
(22) Aleman, C.; Subirana, J. A.; Perez, J. J. *Biopolymers* **1992**, *32*, 621–631.
(23) Kidera, A.; Gō, N. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3718.
(24) Toniolo, C.; Bonora, G. M.; Bavoso, A.; Benedetti, E.; Blasio, B. D.; Pavone, V.; Pedone, C. *Macromolecules* **1986**, *19*, 472–479.
(25) Kitchen, D. B.; Hirata, F.; Kofke, D. A.; Westbrook, J. D.; Levy, R. M.; Kofke, D.; Yarmush, M. *J. Comput. Chem.* **1990**, *11*, 1169.
(26) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
(27) Berendson, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(28) Gō, N.; Noguti, T.; Nishikawa, T. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 3696–3700.
(29) Levitt, M.; Sander, C.; Stern, P. S. *Int. J. Quant. Chem.* **1983**, *10*, 181–199.
(30) Brooks, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6571–6575.
(31) Kitao, A.; Gō, N. *J. Comput. Chem.* **1991**, *12*, 359–368.
(32) Kitao, A.; Hirata, F.; Gō, N. *Chem. Phys.* **1991**, *158*, 447–472.

$$\mathbf{U}^T\mathbf{r}\mathbf{U} = \mathbf{R} \tag{4}$$

where **R** is a diagonal matrix containing the variance of the PC axes, and **U** is normalized as

$$\mathbf{U}_T\mathbf{U} = \mathbf{I} \tag{5}$$

With the average MD conformation placed in an orientation that fits best with the energy minimized conformation, the correspondence between the NM eigenvectors, α_i, and the PC eigenvectors, u_j, is given by their inner products

$$g_{ij} = \alpha_i{}^T u_j \tag{6}$$

A knowledge of the inner products then allows one to express the ith NM vector as a linear combination of the PC axes

$$\alpha_i = \sum_{j=1}^{3N-6} g_{ij}u_j \tag{7}$$

The inner products satisfy the following relation

$$\sum_{j=1}^{3N-6} g_{ij}^2 = 1 \tag{8}$$

and therefore g_{ij}^2 represents the fractional correspondence between the ith NM vector and the jth PC vector.

Although both the NM and the PC axes represent collective motion of the peptide, the fundamental difference lies with the potential energy surface that gives rise to such fluctuations—it is harmonic by definition for the NM, whereas for the PC it may not be so. If the potential energy surface is indeed purely harmonic giving rise to Gaussian distribution in each PC axis, the variance of the individual PC can be used to define angular frequencies that should match that of the NMs according to

$$\omega_j^2 = \frac{k_B T}{R_j} \tag{9}$$

3. Results and Discussion

3.1. Energy Minimization. Energy minimization indicated the α-helix to be lower in energy than the 3_{10}-helix by 0.94 kcal/mol. The average dihedral angles over residues 2–7 were found to be ($\phi = -52.8°$, $\psi = -55.3°$) for the α-helix and ($\phi = -54.9°$, $\psi = -35.4°$) for the 3_{10}-helix. The dihedral angles for the α-helix compare favorably with a previous energy minimzation study[6] and lie well within the experimental spread of ϕ and ψ values for α-helix Aib peptides.[4] The corresponding values of ϕ and ψ for the 3_{10}-helix in the present calculation were found to be smaller than the earlier energy minimization study ($\phi = -49.9°$, $\psi = -31.1°$)[6] and the X-ray crystal structure of a 10-residue Aib homo oligomer ($\phi = -54.1°$, $\psi = -31.2°$).[33] This might be a result of the particular potential energy functions and dielectric constant that we used; energy minimization is not the primary focus of this work, and we will not discuss this point any further here.

3.2. Normal Mode Analysis. Frequency histograms of the calculated NM frequencies for AIB8 in the two helical forms are shown in Figure 1. The histograms are not much different from each other except in the low-frequency region. The 3_{10}-helix possesses a larger number of modes than the α-helix below 300 cm^{-1}. Because the low-frequency modes contain most of the peptide dynamics, qualitatively this suggests the α-helix to be more rigid than the 3_{10}-helix. The rigidity of the two forms will be examined later by quantitative entropy estimations. In Table 1 the ten lowest modes in the two helical forms are shown. Also included in Table 1 are results for the corresponding alanine octamer ALA8 (Ac-Ala$_8$-NHMe). Interestingly, we find that

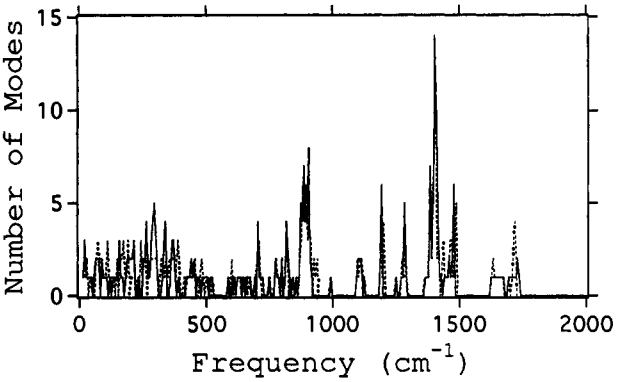

Figure 1. Histograms of calculated normal mode frequency distribution in a (2π) cm^{-1} interval for AIB8 in α-helix (dotted line) and 3_{10}-helix (continuous line).

Table 1. Low-Frequency Modes in the Helices

peptides (helix type)	frequencies of ten lowest normal modes (cm^{-1})
AIB8 (α)	24.1, 32.4, 33.1, 40.2, 51.8, 56.5, 66.9, 70.2, 72.2, 74.6
AIB8 (3_{10})	14.7, 18.3, 19.0, 23.9, 28.3, 34.5, 40.7, 48.7, 60.1, 63.7
ALA8 (α)	20.1, 32.8, 35.0, 39.0, 50.1, 55.0, 65.7, 67.0, 69.0, 71.2
ALA8 (3_{10})	14.5, 18.4, 19.2, 24.9, 28.2, 34.8, 41.2, 47.4, 61.3, 64.8

both the 3_{10}-helices have three softer modes with frequencies lower than the lowest frequencies of the corresponding α-helices.

In Figure 2 motions along the three lowest NMs for both the α- and the 3_{10}-helices are shown. In both the helical forms, most of the motions involve the tangential (twisting) component mixed with bending and stretching, and very often the peptide termini contain most of the displacement. In the 3_{10}-helix mode 1 (14.7 cm^{-1}) involves mostly the twisting of the helix—most of it concentrated at the termini rather than in the middle. Modes 2 (18.3 cm^{-1}) and 3 (19.0 cm^{-1}) are helix-bending modes.

At a qualitative level, the motions of the α-helix did not show any particular characteristic that would make it different from the 3_{10}-helix. Mode 1 (24.1 cm^{-1}, like that in the 3_{10}-helix, shows correlated twisting motion of the two termini. Mode 2 (32.4 cm^{-1}) is a bending motion. In mode 3 (33.1 cm^{-1}) the two termini opened and closed in a concerted manner.

Interhelical conformational entropy difference within the harmonic approximation of the NM,[34] was estimated according to

$$\Delta S_{\text{conf}}^{310-\alpha} = \frac{-k_B}{2} \ln\left(\frac{\det(\mathbf{F}_{310})}{\det(\mathbf{F}_\alpha)}\right) \tag{10}$$

The conformational entropy difference is a quantitative way of elaborating our previous statement that the α-helix is more rigid than the 3_{10}-helix. At 300 K the 3_{10}-helix is found to be entropically favored over the α-helix by 3.75 kcal/mol. Adding this entropic term to the potential energy difference, the free energy of the 3_{10}-helix is lower at 300 K than the α-helix by 2.8 kcal/mol, identical with a previously reported number[7] estimated for the same molecule from potential of mean force calculations using the same force field. Since our entropy calculations assume a simple two-state harmonic model, the agreement could be fortuitous, nonetheless, the importance of the entropic contribution to the 3_{10}-/α-helix equilibrium is clear. Free energy estimates from MD simulation performed in the present work will be discussed later.

NM analysis is, by definition, a harmonic analysis about a minimum. Thus, while NM analysis allows us to identify important collective motions about a particular minimum, it does not shed light directly onto important modes that connect two or more minima. In the next section we will precisely do that by

(33) Toniolo, C.; Crisma, M.; Bonora, G. M.; Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, C.; Santini, A. *Biopolymers* **1991**, *31*, 129–138.

(34) Go, N.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 535–542.

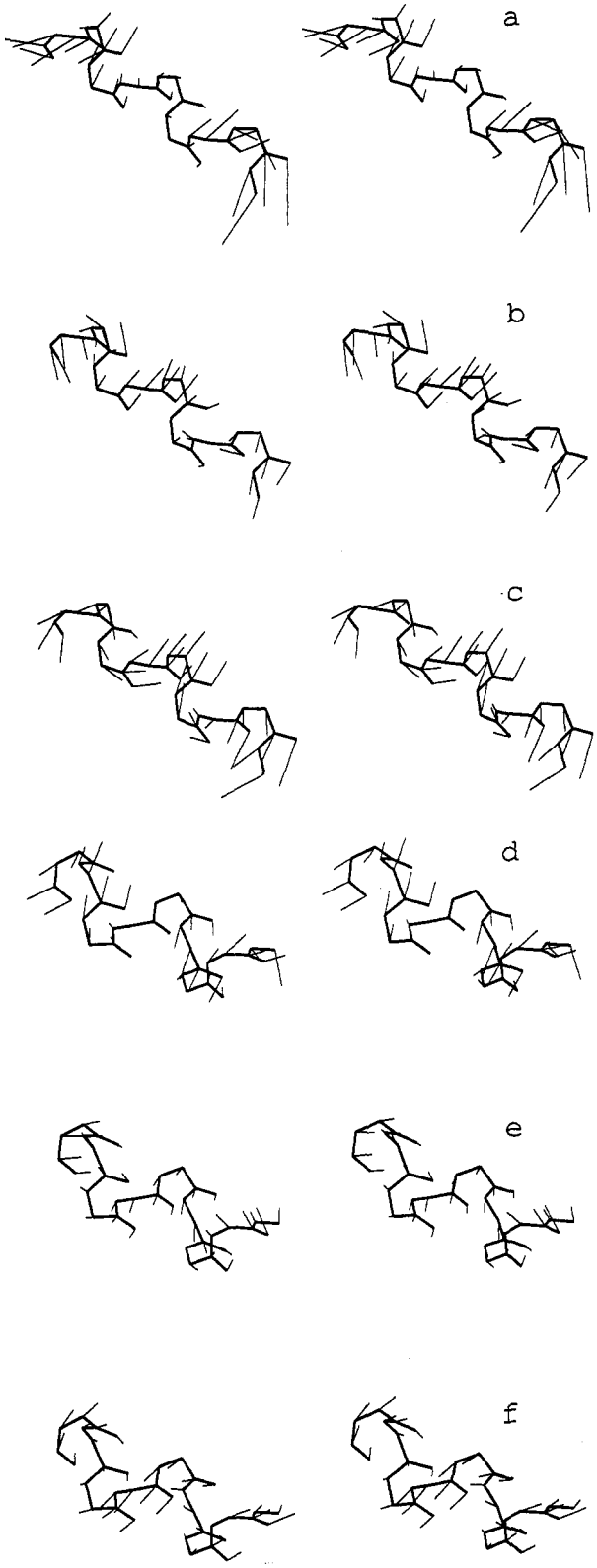

Figure 2. Stereo drawings of atomic displacements in normal modes for the 3_{10}-helix (a) mode 1, (b) mode 2, and (c) mode 3, and, for the α-helix (d) mode 1, (e) mode 2, and (f) mode 3. In all pictures the peptide backbone (thick line) is oriented with the N-terminus at the top left corner. The displacement vectors (thin lines) are shown magnified by a factor of 400 compared to their root-mean-square deviation at the room temperature.

analyzing a long unconstrained MD simulation by the method of PC analysis. Once we identify collective variables spanning different minima we will return briefly to the NM description in terms of projection of the NMs onto the PC axes.

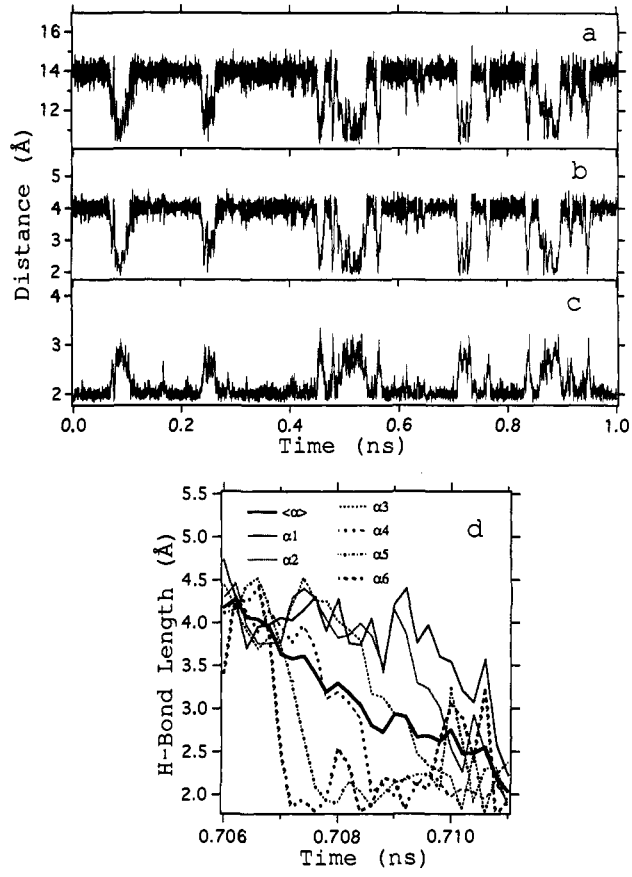

Figure 3. Molecular dynamic simulation at 298 K for the peptide AIB8: (a) the end-to-end length $d^{\alpha\alpha}$, (b) the average CO–HN ($i \rightarrow i + 4$; α-) length $\langle\alpha\rangle$, (c) the average CO–HN ($i \rightarrow i + 3$; 3_{10}-) length $\langle t\rangle$ for the first 1 ns, and, (d) the individual α-H-bond lengths αi (i numbered serially from the N-terminus) from 706 to 711 ps.

3.3. Molecular Dynamics Simulation. The first 1 ns of the 2 ns MD trajectory for AIB8 is shown in Figure 3. Figure 3a shows the end-to-end distance $d^{\alpha\alpha}$ (between the C^{α} atoms of the terminal residues). Figure 3 (parts b and c) show the averages of all possible CO–NH (six α- and seven 3_{10}-type) hydrogen-bond (H-bond) lengths ($\langle\alpha\rangle$ and $\langle t\rangle$, respectively). From Figure 3 (parts b and c) one observes that, while the range of $\langle\alpha\rangle$ fluctuation is about 2 Å (from 2 to 4 Å), $\langle t\rangle$ fluctuation is restricted to about 1 Å (from 2 Å to about 3 Å). The characteristic conformational states encountered with one typical example for each case is given below: (1) 3_{10}-helical regions (from about 0.30 to a little after 0.40 ns) where $\langle t\rangle$ remains at 2.0 Å and $\langle\alpha\rangle$ is about 4.0 Å; (2) Mixed 3_{10}-/α-helical regions (around about 0.25 ns) where $\langle t\rangle$ and $\langle\alpha\rangle$ fluctuates around 2.5 Å; and (3) α-helical regions (around 0.73 ns) where $\langle t\rangle$ fluctuates between 3.0 and 2.5 Å and $\langle\alpha\rangle$ fluctuates between 2.0 and 2.5 Å.

The end-to-end distance, $d^{\alpha\alpha}$, shown in Figure 3a, also changes concomitantly along with the changes in the helical states. In the 3_{10}-helix, $d^{\alpha\alpha}$ fluctuates between 13.0 and 15.0 Å, in the mixed helical states (around about 0.25 ns) this fluctuation is restricted between 11.0 and 13.0 Å, while in the α-helix $d^{\alpha\alpha}$ fluctuates between 10.0 and 11.0 Å. In the next 1ns of the MD run similar features were observed.

One typical case of the transition ($3_{10} \rightarrow \alpha$; just after 700 ps) is shown in Figure 3d to demonstrate an interesting mechanism of the transition that was encountered several times in the run. Although during the transition the average of all possible six α-H-bond length changes smoothly from about 4 to 2 Å, we show the individual α-H-bond lengths (α1, α2, etc. from the N-terminus) in order to take a closer look into the transition region. The principal observation is that of a boundary propagation

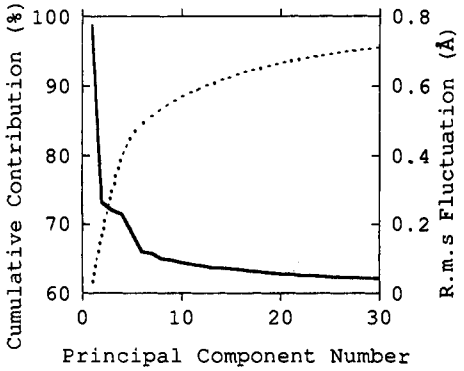

Figure 4. Root-mean-square fluctuation (solid line) and the cumulative contribution (broken line) to the total fluctuation from the first 30 principal components.

mechanism for the transition. The transition starts at the C-terminus and propagates towards the N-terminus in a sequential manner. In other instances of transition the opposite direction of propagation was also encountered. Also, after the transition, when the peptide maintains more-or-less the α-helical form, only $\alpha 3$, $\alpha 4$, $\alpha 5$ and to some extent $\alpha 6$ remain close to 2 Å while $\alpha 1$ and $\alpha 2$ almost always remain broken. A look at the corresponding 3_{10}-H-bond lengths (not shown here) showed that t1 remained close to 2 Å during this time, t2 occasionally becomes closer to 2 Å, while all the rest remained broken. These mixed helical states appeared frequently during the run.

To summarize, the MD data showed transitions between the two helical states with the obvious involvement of the mixed helix indicating that a simple assumption of a two-state problem, that of a *pure* 3_{10}- and a *pure* α-helix is not valid. Within the α-well the peptide visited several conformations—that of mostly α- with a mixed 3_{10}-form. The mixed helical states were more accessible from the α-well than the 3_{10}-well. The transitions, of which one typical case was described above, were not concerted throughout the helix but were sequentially propagational. In a less typical case we also observed a mechanism where propagation occurs with two contiguous H-bonds as a unit.

3.3.1. Principal Component Analysis. To identify important collective motions during the 3_{10}-/α-helix transition, we next analyzed the 2 ns MD trajectory by the method of PC analysis. This analysis has been used in the past to analyze peptide fragments in the protein data bank,[35] to refine structure from X-ray crystallography,[23] and to analyze MD trajectories.[32,36]

After removing any rotational or translational motions associated with the conformations generated by the MD run, the variance–covariance matrix was diagonalized to obtain the PCs as described earlier. The root-mean-square fluctuation (rmsf) along each PC axis and the relative contribution of each PC to the mean-square fluctuation is shown in Figure 4. The percent contribution toward the total fluctuation of the first ten PCs are 61.5, 7.1, 5.9, 5.3, 3.1, 1.4, 1.3, 1.0, 0.9, and 0.8 respectively; cumulatively the first ten PCs account for about 88.3% of the total fluctuation. We will therefore concentrate on the first ten PCs hereafter. Clearly fluctuations described by the first few PCs dominate the total fluctuation of the peptide, and the first PC axis, PC1, by itself represents almost two-thirds of the total fluctuation. This fact, that a single collective variable, PC1, accounts for a major portion of the dynamics out of a total of 342 internal variables, singles out PC1 as a vector along which the transition process should be examined.

One convenient way to understand the conformational changes that occurred during the MD run is to project the MD trajectory onto the PC axes. This was done by expressing each PC coordinate

(35) Takahashi, K.; Gō, N. *Biophys. Chem.* **1993**, *47*, 163–178.
(36) Hayward, S.; Kitao, A.; Hirata, F.; Gō, N. *J. Mol. Biol.* **1993**, *234*, 1207–1217.

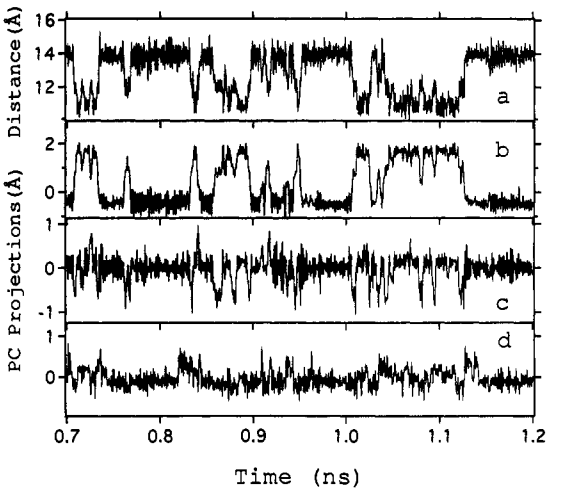

Figure 5. (a) The end-to-end distance, $d^{\alpha\alpha}$, and projections on (b) PC1 axis, (c) PC2 axis, and (d) PC3 axis, from 700 to 1200 ps in the MD simulation.

as a linear combination of the mass-weighted displacements of atoms in the MD trajectory as

$$d_j(t_a) = \sum_{i=1}^{3N} \Delta X_i(t_a) u_{i,j} \qquad (11)$$

where $d_j(t_a)$ represents projection of the t_ath conformation in the MD trajectory onto the jth PC axis. These projections for the first three PC axes are shown in Figure 5 along with $d^{\alpha\alpha}$ fluctuation. As pointed out earlier in Figure 3a, $d^{\alpha\alpha}$ correlates well with the helical state of the peptide. In Figure 5 good correlation is also observed between $d^{\alpha\alpha}$ and projection onto the PC1 axis—they appear as almost perfect mirror images of each other. This fact makes the PC1 axis relevant and unique in relation to the 3_{10}-/α-helix transition. Unlike the case with PC1 axis, projections on PC2 and PC3 axes do not show straightforward correlation with $d^{\alpha\alpha}$. Upon a closer examination, however, one can see that d_2 deviates sharply from its mean, either in the positive or negative direction, whenever there is a transition or whenever the peptide is in a mixed helical state. This picture will be clearer in Figure 6 where the trajectory information is projected in two-dimensional conformational subspaces spanned by the lowest five PC axes.

Projection of the MD trajectory onto two PC axes i and j results in a two-dimensional probability distribution $P(d_i,d_j)$. From this probability distribution one can estimate the free energy difference between any two points on the two-dimensional plane spanned by the ith and the jth PC. This free energy difference is given by

$$\Delta F(d_i,d_j) = -RT \ln\left(\frac{P(d_i,d_j)}{P(3_{10})}\right) \qquad (12)$$

where the absolute free energy of the 3_{10}-helix is set to zero. Contour maps of free energy surfaces spanned by d_1 and the other four lowest PC coordinates, d_2, d_3, d_4, and d_5, are shown in Figure 6 along with the visually traced most populated pathways of transition as encountered in the MD run. Figure 6 captures the essential dynamics of the MD run. Apart from the two clear energy minima (3_{10}-helix, d_1 = 0.5 Å; α-helix, d_1 = 1.75 Å) on the PC1/PC2 plane (Figure 6a), two distinct pathways of interconversion are conspicuous—one takes the negative d_2 route (the down pathway), while the other takes a positive route (the up pathway). The down pathway is more populated. On the PC1/PC3 plane two conspicuous minima are still visible. The two minima are connected by a maximum populated pathway,

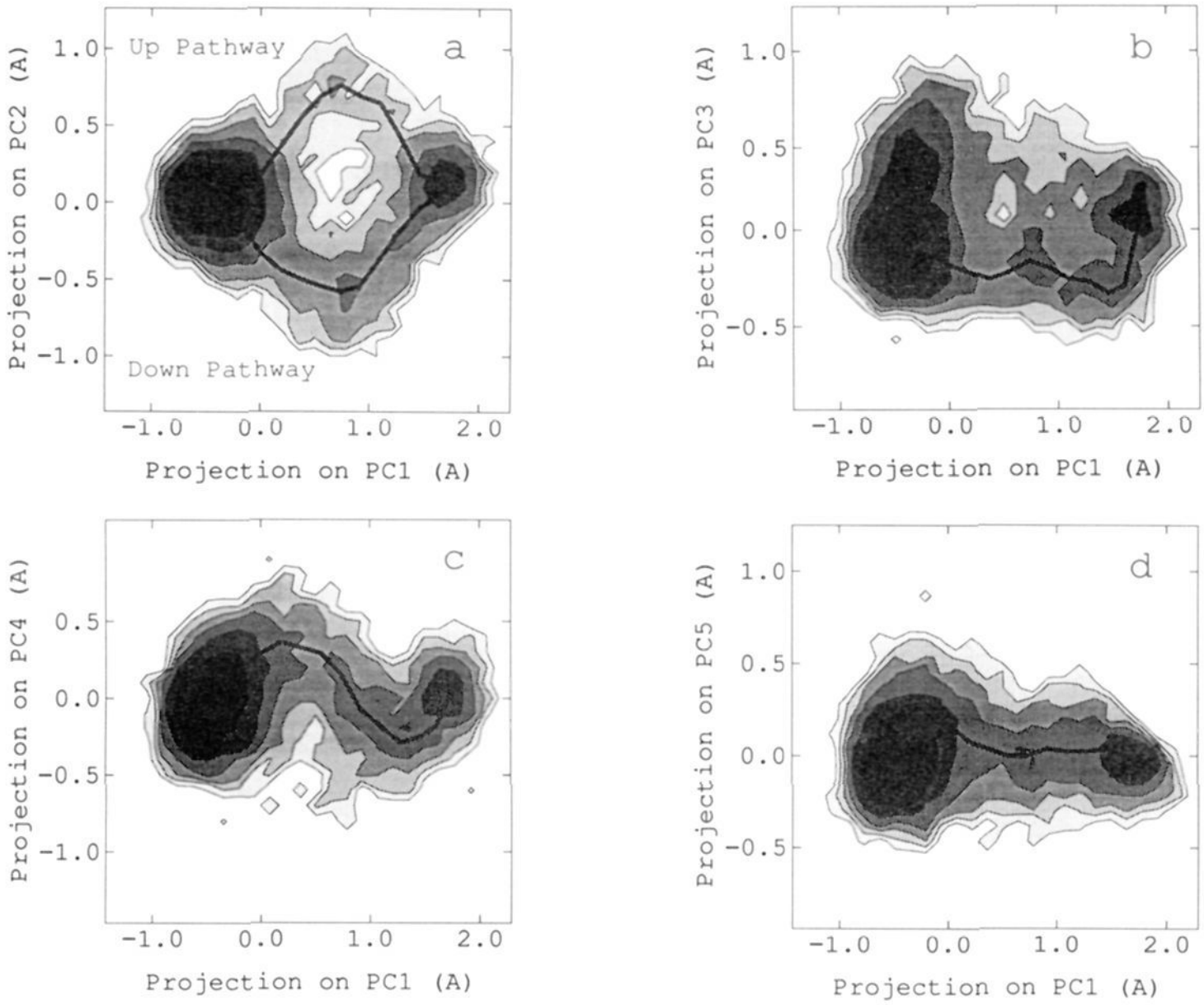

Figure 6. Contour maps of free energy surfaces spanned by (a) PC1/PC2, (b) PC1/PC3, (c) PC1/PC4, and (d) PC1/PC5 axes. The maximum populated pathways for transition are indicated by thick solid lines. The contour lines are drawn every 0.5 kcal/mol with the darker shades indicating lower free energy values. Two pathways are identified as the up and the down pathway in Figure 6a.

the width of which is broader than in Figure 6a. On the PC1/PC4 and PC1/PC5 planes (Figure 6c,d) free energy variations are also simpler in the sense that they do not suggest any obvious multipathway mechanism. We find that d_4 undergoes a "twisty" nonmonotonic change, while d_5 changes more-or-less monotonically. Understanding the transition in terms of the free energy variations along these first five PC axes will help us to construct some plausible reaction pathways.

Apart from identifying pathways, the free energy contour maps in Figure 6 also allow direct estimation of interhelical free energy differences and activation free energies. Both the interhelical free energy difference, $\Delta F(\alpha\text{-}3_{10})$, and the activation free energy, $\Delta F^{\ddagger}(\alpha \rightarrow 3_{10})$, are about 1 kcal/mol. The NM estimate for $\Delta F(\alpha\text{-}3_{10})$ was higher by about 1.8 kcal/mol than the direct MD estimate. Differential anharmonicity effects in the two helical wells may give rise to this difference. Small activation free energies were also found by Marshall and co-workers for the 3_{10}-/α-helix transition in Aib nonamer[14] where the α-helix, rather than the 3_{10}-helix, was found to be the stabler form, contrary to experimental suggestions.[33] The characteristic time scale of interhelical transition in AIB8 in vacuum is 10^{-10} s, estimated directly from the MD trajectory. However, inclusion of solvent effects is essential for any realistic rate estimation.

Motions along any PC axes describe collective motions just as do the NMs in the harmonic analysis. Motions along the first three PC axes are shown in Figure 7. A twist of the helix characterizes motion along the PC1 axis (rmsf 0.77 Å). In fact, when viewed down the helical axis, the peptide backbone transformed from the 3_{10}-form (eclipsed triangles) to the α-form (staggered triangles) as shown in Figure 8. The helix is disrupted/stabilized alternatively in the two halves of the peptide along the PC2 axis (rmsf 0.26 Å). Unwinding of the N-terminus along with a slight twist at the C-terminal half characterizes motion along the PC3 axis (rmsf 0.24 Å).

3.3.2. Correspondence between NM and PC Axes. It will be instructive at this point to examine the correspondence between the NM axes defined at each helical minimum and the PC axes defined by averaging over the MD trajectory. As discussed earlier, this was achieved by calculating the inner products between the PC eigenvectors and the NM eigenvectors (eq 7). The inner products allow each PC axes to be expressed as a linear contribution of the NM axes, and the square of the inner products between two eigenvectors reflect the fractional correspondence between the two. In Table 2 we list the square of inner products between the first five PC and NM axes. The correspondence between PC1 axis and NM1 axes in the two helical minima is 64%. The second PC axis is represented in the 3_{10}-minimum by the NM2 (45%) and NM4 (42%) axes and in the α-minimum by the NM3 (81%) axis. The cumulative correspondence between the first five axes of NM and PC is also high (from 46% to 94%). In Figure 9 we show the correspondence between NM and PC axes with frequencies (effective frequency for PC axes calculated according to eq 9) up to 300 cm^{-1}. Correspondence of more than 4% is indicated by a dot while a filled circle indicates more than 20% correspondence. The corresponding effective frequencies of the PC axes were found to be shifted toward lower values. Since the conformational space spanned by the PCs include both the α- and the 3_{10}-minima, they are characterized by a slightly larger mean-square variance, and this causes the shift.

The first few PC axes define a conformational subspace that can be considered the most important, because they contain most of the total fluctuation. This important subspace contains both the α-helix and the 3_{10}-helix conformations. In a similar fashion, one can envision subspaces spanned by a set of lowest frequency

Table 2. Square of Inner Products, $g_{i,j}^2$, between NM and PC Axes

PC axes	NM axes (3_{10}-minimum)						NM axes (α-minimum)					
	1	2	3	4	5	ΣNM	1	2	3	4	5	ΣNM
1	0.64	0.00	0.01	0.00	0.01	0.66	0.64	0.01	0.01	0.02	0.00	0.68
2	0.00	0.45	0.01	0.42	0.00	0.88	0.02	0.01	0.81	0.00	0.01	0.85
3	0.00	0.21	0.00	0.21	0.14	0.56	0.00	0.12	0.02	0.41	0.00	0.55
4	0.14	0.01	0.72	0.04	0.03	0.94	0.07	0.67	0.00	0.08	0.02	0.84
5	0.00	0.25	0.06	0.13	0.32	0.76	0.01	0.04	0.06	0.01	0.34	0.46
ΣPC	0.78	0.92	0.80	0.80	0.50		0.74	0.85	0.90	0.52	0.37	

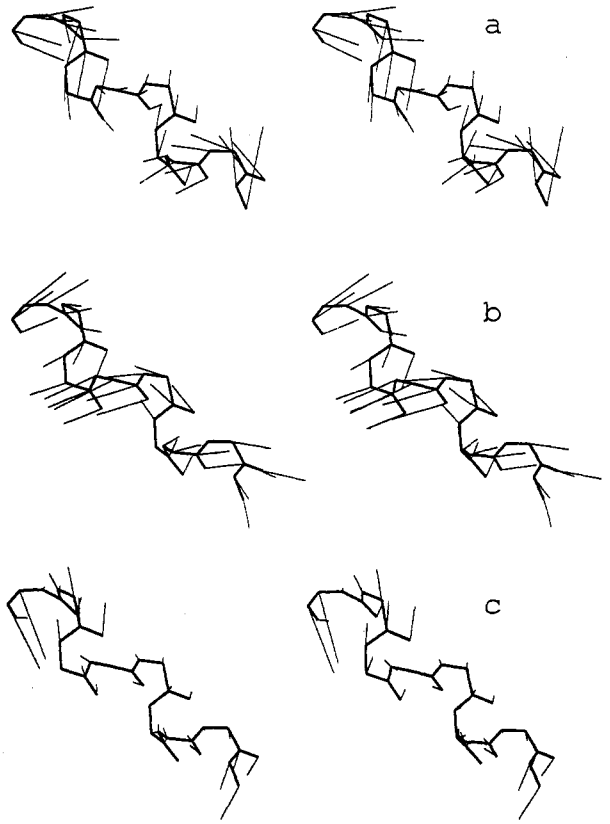

Figure 7. Stereo drawings of atomic displacements in principal component vectors (a) 1, (b) 2, and (c) 3. The displacement vectors are shown magnified by a factor of 100 for vector 1 and 300 for the vectors 2 and 3 compared to their rmsd.

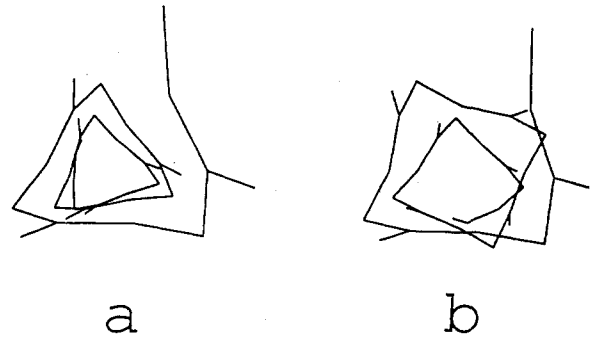

Figure 8. Two snapshots from peptide motion in principal component vector 1 (Figure 7a) viewed from the N-terminus down the helical axis. The PC1 coordinates for the two structures correspond to (a) –0.5 Å (3_{10}-helix), and (b) 1.75 Å (α-helix).

NM axes in both the helical minima. However, because these NM subspaces are only restricted to one type of helix, they may not be very realistic in capturing the most important part of the total conformational space. Correspondence between the PC subspace and the NM subspaces thus become an important test for judging the relevance of the NM subspace in describing fluctuations in a realistic fashion. For the present case, we indeed

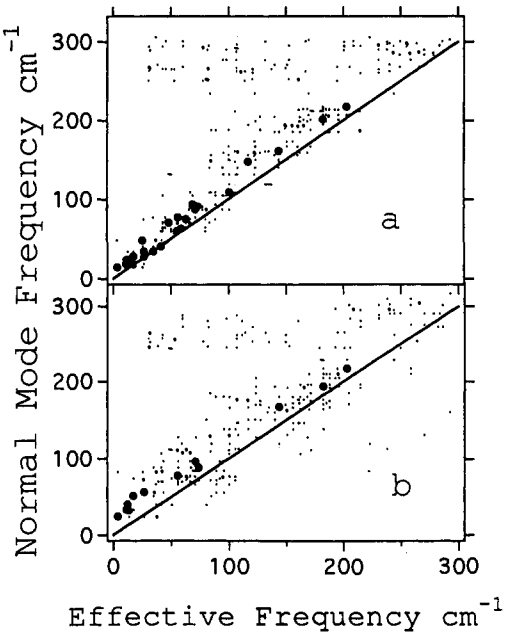

Figure 9. Inner product squares (eq 6) between (a) the 3_{10}-helix NM vectors and the PC vectors and (b) the α-helix NM vectors and the PC vectors. Values of g_{ij}^2 greater than 0.04 are indicated by a dot and those greater than 0.2 are indicated by a filled circle.

find a high degree of correspondence, and this endorses the concept of important subspace, that is, subspaces spanned by low-frequency NM axes can represent a major share of the total fluctuation observed on energy surfaces not necessarily harmonic.[37]

3.3.3. Transition at a Molecular Level. Having identified the important subspace spanned by PCs that account for a major share of the total fluctuation and those that are most important in contributing toward the 3_{10}-/α-helix transition, we now turn into constructing pathways of transition in this subspace. We had earlier observed in the raw MD data that transitions were not concerted in terms of breaking or making the 3_{10}- ($i \rightarrow i+3$) and the α-($i \rightarrow i+4$) H-bonds (Figure 3d). We will examine if the pathways constructed in a given subspace can yield the right mechanism.

Because PC1 is now known to be most responsible for the transition, we at first construct a pathway in the one-dimensional space spanned by PC1 only. The conformational change along this pathway is shown in Figure 10a, where the individual α-H-bond lengths are plotted along this pathway, i.e., as a function of d_1 from about –0.5 Å for the 3_{10}-helix to about 1.75 Å for the α-helix. We see that this one-dimensional pathway is sufficient to bring about the 3_{10}-/α-helix transition. The mechanism is a concerted one and does not match with what we observed from the raw MD data (Figure 3d).

While the first PC axis allows the peptide to move from one helical form to the other, one can expect the next few PC axes to add the "right mechanism" to this transition. In Figure 10b–f we show the behavior of the individual α-H-bond lengths, when pathways are constructed in the two-dimensional space involving

(37) Hayward, S.; Kitao, A.; Gō, N. *Protein Science*, in press.

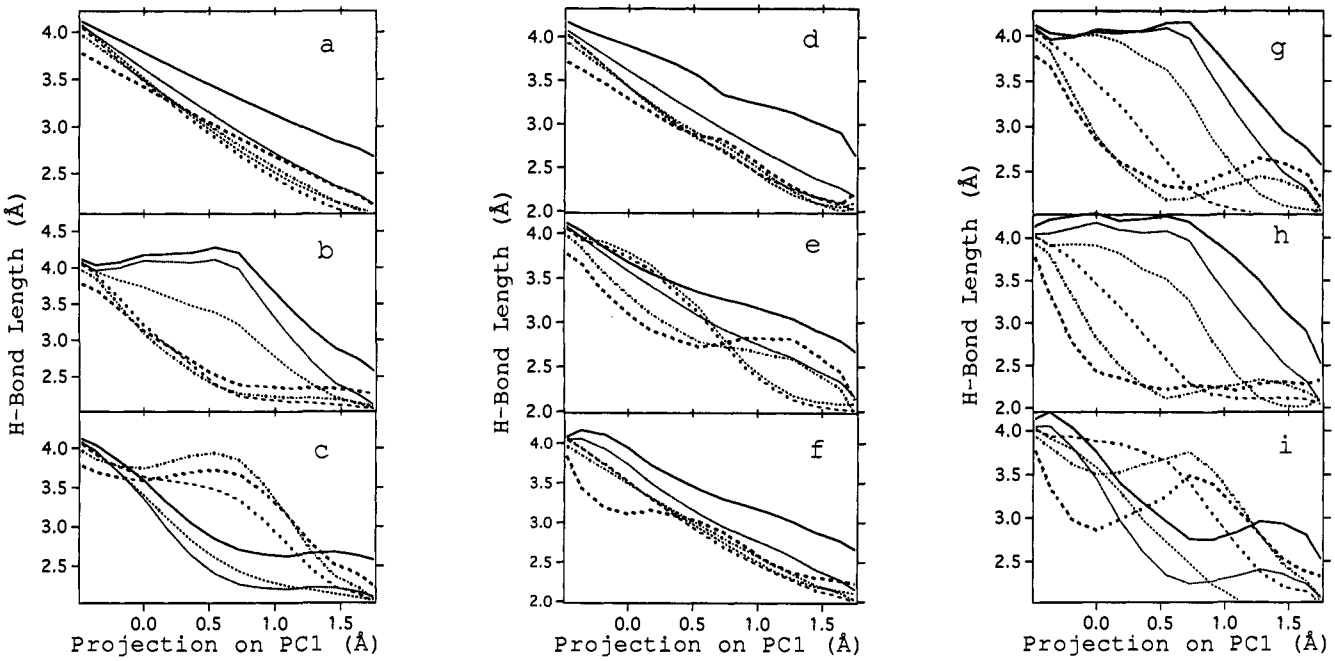

Figure 10. Individual α-H-bond lengths αi (i numbered serially from the N-terminus) are shown as a function of the first PC coordinate. Line notations are the same as in Figure 3d. The reaction pathways span a one-dimensional space in (a) PC1, a two-dimensional space in (b) (PC1 and PC2-down), (c) PC1 and PC2-up, (d) PC1 and PC3, (e) PC1 and PC4, and (f) PC1 and PC5, a three-dimensional space in (g) PC1, PC2 down, and PC4, and a five-dimensional space in (h) PC1, PC2-down, PC3, PC4 and PC5 and (i) PC1, PC2-up, PC3, PC4 and PC5.

PC2, PC3, PC4, or PC5, together with PC1. In the space spanned by PC1 and PC2, we considered two pathways, up and down, shown by thick solid lines in Figure 6a. In the PC1/PC2 space, the peptide behaves like two parts joined together. Transition in one-half precedes that in the other half, and depending on if one takes the down pathway (Figure 10b) or the up pathway (Figure 10c), the $3_{10} \rightarrow \alpha$ transition is initiated from the C-terminal half or the N-terminal half, respectively. When the contribution from PC3 is added to contribution from PC1 (Figure 10d), there is not much change of the mechanism from Figure 10a except for the behavior of the α1 H-bond length. Combined contributions from PC1 and PC4 on the other hand adds an interesting feature to the mechanism (Figure 10e) arising from PC4. In the first half of the transition α3 and α4 H-bond lengths become longer while α5 and α6 H-bond lengths become shorter and vice versa in the second half of the transition. Contribution from PC5, like that from PC3, is also found to be minimal in affecting the transition mechanism (Figure 10f). To summarize, of the first five PC axes, PC1 enables the transition, PC2 and PC4 add features that are crucial to the mechanism, while contributions from PC3 and PC5 toward the transition are minimal. How realistic a mechanism can one generate if one includes these three PC axes—PC1, PC2, and PC4, only? This is shown in Figure 10g. This picture is to be compared with Figure 3d. Despite the fact that Figure 3d contains contributions from all degrees of conformational freedom, and Figure 10g includes contributions from only three degrees of freedom, the similarities between the two figures are striking. Both figures suggest a transition with boundary propagation. In Figure 3d we observed a complete sequential mechanism, and Figure 10g suggests the same except for the behavior of the α5 and α6 H-bond lengths, which change in a concerted way. Essentially we have succeeded in reducing the essential degrees of freedom for the transition from 342 to 3, and now we examine the effect of considering all the first five PC axes.

When considering all the first five PC axes, we construct two pathways arising from the fact that transitions projected on the PC1/PC2 axes exhibited a bifurcated nature (Figure 6a). The conformational changes in the two are shown in Figure 10h,i. The more populated of the two, the down pathway, exhibits the sequential mechanism (Figure 10h), and the less populated up pathway exhibits a semisequential mechanism (Figure 10i), where pairs of contiguous α-H bonds are formed in a concerted manner. Both these mechanisms in fact were found during the MD run. Also, Figure 10h, which differs from Figure 10g in that Figure 10h includes additional contributions from PC3 and PC4, α5 and α6 H-bond formation is completely sequential. The combined combination from all the first five PC axes thus recreates Figure 3d almost completely. This justifies the fact that we restricted ourselves to only the first five PC axes in describing the transition.

A concerted mechanism of the 3_{10}-/α-helix transition, published in an earlier work,[6] is definitely not a valid mechanism. The most populated pathway is a sequential boundary propagation mechanism where the $3_{10} \rightarrow \alpha$-helix transition is initiated from the C-terminus. In a recent study sequential mechanism has been identified as the low-energy pathway for the transition in a longer Aib peptide.[14] Apart from the sequential mechanism, we also found that other mechanisms can be operative, namely where contiguous pairs of α-H-bonds form or break. In addition, our work is different from previous studies of transition pathways in that we identified the most important degrees of collective motion responsible for the transition. In the present case only three were shown to contribute significantly during the transition. Identifying such collective models can be of further significance. Since the first PC axis alone was found to be able to account for the transition, PC1 axis can be used as an excellent reaction coordinate for this transition, along which further theoretical investigations can be undertaken. For example, analysis of solvent induced potential of mean force along the reaction coordinate based on either molecular simulations or the integral equation method will provide a realistic view for the peptide transition in solution.[38]

It is to be noted that when the duration of transitions is extremely short relative to the interval between them (case of very quick transition), only one PC, usually PC1, a vector connecting the two states before and after transition, contains information about the transition. In such a case, PC analysis will produce only a first-order description of the pathway in terms of PC1. The system

(38) Kitao, A.; Hirata, F.; Gō, N. *J. Phys. Chem.* **1993**, *97*, 10231–10235.

we studied spends a significant amount of time in the transition region, so that better than a first-order description of the pathway could be attained in terms of a few more principal components.

4. Conclusion

The 3_{10}-/α-helix equilibrium in Aib containing peptides is a very special transition, yet, because of the conformationally narrow helical wells involved, this is an ideal case to study peptide dynamics from computer simulations of relatively short time. For the peptide AIB8, an eight-residue Aib homooligomer, we found that potential energy minimization alone was insufficient to predict the experimental 3_{10}-helical preference of the peptide. Conformational entropy estimates from a harmonic analysis reversed the theoretical helical preference with the 3_{10}-helix lower than the α-helix by 2.8 kcal/mol. NM analysis also revealed the 3_{10}-helix to be more flexible than the corresponding α-helix. Especially the 3_{10}-helix possessed three modes below 24 cm^{-1}, while the lowest mode for the α-helix began beyond this frequency. Thus the interhelical entropy difference affects the helical equilibrium significantly.

Free energy differences were also estimated directly from MD run. The NM method was found to overestimate the interhelical free energy difference when compared to the MD estimate—about 1 kcal/mol. Free energies of activation for the $\alpha \rightarrow 3_{10}$ transition was about 1 kcal/mol and that for the $3_{10} \rightarrow \alpha$ transition was about 2 kcal/mol.

Exploring beyond the harmonic model of collective peptide dynamics, we considered a long MD simulation and analyzed the trajectory by the method of PC analysis—a quasi-harmonic method. Our focus was to identify collective motions of the peptide, as was done in the NM analysis, except now the collective modes are not restricted to a single helical minimum. The first few collective modes (PC) were shown to be dominantly important in defining peptide dynamics during the transition. Fluctuations along the first PC axis, which represents about two-thirds of the total fluctuation, enables the peptide to traverse both the helical minima. When fluctuations along only two other PC axes (2 and 4) were added to this, we reconstructed almost in entirety the molecular mechanism of the transition as observed during the MD run—a sequential boundary propagation ($3_{10} \rightarrow \alpha$) mechanism from the C-terminus. This fact, that only three collective modes could describe the essential features of the transition, is a very significant finding. Identifying the important collective modes responsible for the transition also allowed us to consider mechanisms other than a sequential one, and we found a semisequential mechanism also operative during the transition where two contiguous α-hydrogen bonds form a break in a locally concerted way. A sequential mechanism of the transition has been reported earlier[14] for a longer Aib peptide, but this is the first time that a picture of the transition has been constructed as a superposition of collective modes.

A collective motion picture of the transition has several advantages in itself and can provide important guidelines for further simulations. A transition mechanism takes an added meaning when defined as a superposition of collective modes since one can deconstruct the mechanism by removing layers of collective modes, each of which represents an orthogonal axis in the conformational space. Free energy simulations, frequently performed to investigate helical transitions in peptides, need a predefined reaction coordinate along which further simulations are performed.[1] Very often the reaction coordinate is chosen arbitrarily to be the peptide length[14] or as a function of one of the dihedral angles where the angles vary in tandem.[17] This arbitrariness can be removed if one chooses, in the present case for example, the reaction coordinate to be the PC1 axis—fluctuations along which was shown to induce the 3_{10}-/α-helix transition.

Since we showed that the lowest few PC axes were sufficient to describe the essential transition dynamics, further simulations in a conformational subspace spanned by say the first few PC axes, an idea also put forward in a recent work,[39] can be very rewarding. The idea of molecular simulations along collective modes is not new, Monte Carlo simulation along the NM axes have been performed.[40] Therefore, the PC axes not only are able to define a natural reaction coordinate, but also can be utilized to reduce the dimensionality of the problem.

Acknowledgment. We would like to thank Dr. Steven Hayward and Professor Atsuo Kuki for insightful discussions during the course of this work. We thank Professors Garland Marshall and Jan Hermans for making several preprints available to us. This work was supported by grants from the Ministry of Education, Science and Culture, Japan, and from the International Human Frontier Science Program Organization (HFSPO). G.B. is a recipient of the long-term postdoctoral fellowship from HFSPO (LT/335-92).

(39) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins* **1993**, *17*, 412–425.
(40) Noguti, T.; Gō, N. *Biolpolymers* **1985**, *24*, 527–546.